

Weekly Report

Pingping Shang

2013.9.9~2013.9.15

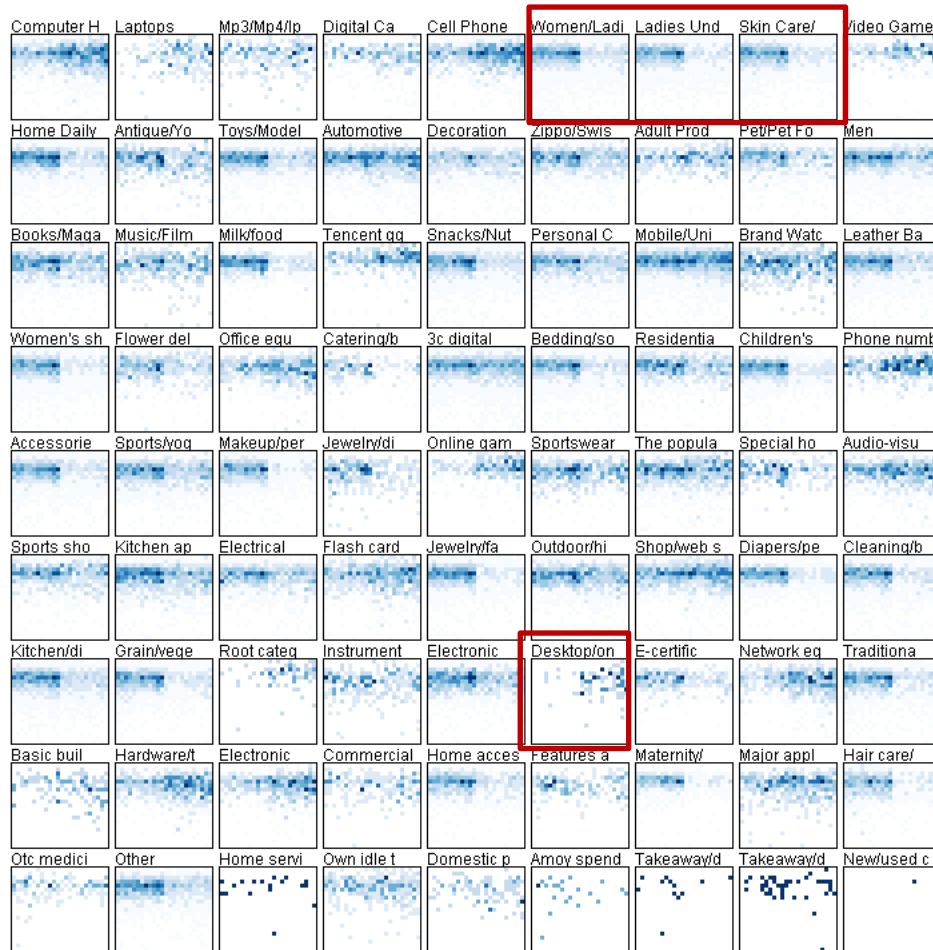
Work in this week

1. Mosaic Project

- 1) Refactor. One problem encountered in this process is how to make the dynamic representation: the area each block occupied in different hierarchies is the same or it will become smaller and smaller with the amount of labels selected increasing. Finally, we applied the second scheme because of the beauty of the interface. Another detail I did not notice is the exchange strategy of location which make it convenient for the analyst to compare two goods or genres.

- 2) Test the effect of Taobao data when make the similar set with Xinxin: Gender and Constellation are listed in the horizontal direction and Age is listed in the vertical direction. The effect is presented below.

We can find the obvious boundary line in horizontal direction in some blocks. However, it becomes difficult to present exquisite result as Xinxin's.



2. Read papers

“Mapping Nominal Values to Numbers for Effective Visualization”

1) Problem: Most existing visual exploration tools are designed to handle numeric variables only and techniques that map nominal (or categorical) values to numbers are rather simplistic.

2) New Method: Display nominal variables in visual exploration tools designed for numeric variables.

Summarize with my own words: The paper pays most attention on nominal variables with high cardinality, e.g. product codes. The technique relative with us is investigated to reduce the number of distinct values to display (Distance-Quantification-Classing). Distance step applied Correspondence Analysis method to extract a reduced set of nominal variables that are independent of each other. Classing step will combine the distinct value of a nominal variable and assign space according to the distance between each other.

Plan for next week

Go on to refine the code of Mosaic Project and try the new idea that selects a random area to find the similar group.

Survey papers about categorical data.